Image Super-Resolution of Medical Images

Mireille El Gheche and Pascal Frossard

1 Summary

Recent developments in medical image processing bring the need to solve optimization problems with increasingly large sizes, stretching traditional techniques to their limits. New optimization algorithms need to be designed, paying attention to computational complexity, scalability, and robustness. Proximal approaches have become increasingly popular recently, in both signal/image processing and machine learning areas. The proximal framework provides simple, elegant and flexible way to construct optimization algorithms that benefit from solid theoretical foundations and show great practical efficiency. This project aims at proposing a new generation of proximal algorithms that remain efficient in the context of big data processing, thanks to the integration of parallel, distributed, and online computing strategies.

Among the several applications in image processing, the one that motivates our project is the super-resolution of histological images under the assumption that high-resolution snapshots of the tissue and its corresponding phenotypes are available (Figure 1). Due to the time constraint related to the technology used, the scanner produces a low-resolution image of the whole slides, leading to a loss of important details for the estimation of cell population. Fortunately, the resolution is improved by zooming into specific parts of the sample. A set of high-resolution snapshots provided whose locations are predefined by doctors. These high-resolution snapshots are fed into an analysis software, such as InForm [], which can provide a powerful tool to unveil the phenotype of the cells. However, the number of the high-resolution snapshots are limited with regard to the whole slide size, which may lead to miss some important information in the tissue. Even worse, the snapshots are often chosen on some specific areas with the same properties that makes hard to interpret in terms of the whole slide analysis process.

The goal of this project is to develop a data framework to built new algorithms which are able to produce high resolution images by taking in account the high resolution snapshots and the underlying image structure.

Figure 1: Example of a low resolution medical image with six high-resolution snapshots. Middle: low-resolution of the whole slide with 45 snapshot positions (small boxes). Left and Right: example of six high-resolution snapshots.

he proposed project defines very important research challenges both from the domain application and the optimization perspective.

2 Research plan

This project aims at super-resolving a low-resolution texture under the assumption that high-resolution snapshots of the texture are available. To attain this objective, one needs to define the mathematical characteristic of the underlying texture, and then formulate an optimization problem that combines the data fidelity term and priors of the sought texture. In what follows, we summarize the state-of-the-art related to this proposal and then detail the proposed research plan.

State of the art The simplest way to upscale a low-resolution image is by mean of linear scaling (such as bi-cubic interpolation) or image sharpening methods. However, these methods introduce high frequency concealing, which blur the image and erase small texture details.

A number of single-image reconstruction methods are based on the assumption that a dataset of high-resolution textures are available, from which they learn a dictionary of low-resolution patches, a dictionary of high-resolution patches, and a correspondence map between the two dictionaries [1]. However, the results depend on the estimated dictionaries, and thus the reconstruction of true details is not guaranteed if the target texture does not appear in the training dataset. A different approach consists of introducing a suitable regularization into the reconstruction problem formulation, with the aim of conveying some prior knowledge about the signal to be recovered. In this context, Total Variation [2] its higher-degree directional derivatives with non-locality principle [3] have emerged as regularization, consisting in penalizing the gradient coefficients. However, these regularization terms fail to preserve textures, details, and structures, because they are hardly distinguishable from noise.

Note that this problem can be also related to texture synthesis from a high-resolution patch, guided by a low-resolution texture. Texture synthesis techniques [4] can be broadly categorized into region-growing local methods and optimization-based global methods. Local approaches grow the texture one pixel (or patch) at a time, while maintaining the spatial coherence with nearby pixels by modeling the neighborhoods with Markov fields and fractal models [5]. A weakness of these methods is that the spatial coherence between pixels is enforced at a local scale. Global methods process the entire texture as a whole, using some criteria for measuring its similarity with a small texture patch. For example, the latter can be modeled with a statistical descriptor based on histograms [6], wavelet coefficients [7], or Fourier coefficients [8]. A similar approach was proposed in [9], which introduces a preliminary step of dictionary learning for exploiting the given patch, and (not least) the Wasserstein distance for comparing the histograms of the entire texture with an extended version of the small patch [10].

The methods described so far are unsupervised. Recently, many methods have been proposed to learn from examples how to upsample the textures. However, one needs a sufficient amount of data, pairs of low-resolution and high-resolution textures. In addition, the super-resolution algorithms are overfitted to the training data and unlikely to be generalized.

The problem of interest Let's denote the high-resolution signal of interest $\bar{x} \in \mathbb{R}^N$, which generally corresponds to an image of size $N = N_1 \times N_2$. The degradation model that we consider is the following:

$$z^{(1)} = \mathrm{D}\mathrm{B}\bar{x} + \eta_1 \tag{1}$$

Hereabove, $z^{(1)} \in \mathbb{R}^Q$ is the complete low-resolution image, $\mathrm{B} \in \mathbb{R}^{N \times N}$ is a linear operator modeling some blur, $\mathrm{D} \in \mathbb{R}^{Q \times N}$ stands for spatial down-sampling by a dyadic factor in each direction, yielding $Q = 2^{-r}N$, and

$\eta_1 \in \mathbb{R}^Q$ is a realization of an additive zero-mean white Gaussian noise with standard deviation τ. Moreover, let $\forall i \in \{1, \ldots, I\}$, $z_i^{(2)} \in \mathbb{R}^M$ denotes a small snapshot of the high-resolution image $\bar{x}$, and $z^{(2)} = \{z_1^{(2)}, \ldots, z_I^{(2)}\}$ be the set of the high-resolution snapshots. We propose to recover $\bar{x}$ from the observations $z^{(1)}$ and $z^{(2)}$ through a variational approach that leads to solving the following optimization problem:

$$\underset{x \in \mathbb{R}^N}{\text{minimize}} \ \|\mathrm{D}\mathrm{B}x - z^{(1)}\|^2 + \sum_{j}^{J} \lambda_j R_i(x, z^{(2)}) \tag{2}$$

where $\forall j \in \{1, \ldots, J\}$, $\lambda_j > 0$ is a regularization parameter. Beside the data fidelity term w.r.t. the observations $z^{(1)}$, we will use different additional piece of information $R_j(\cdot, z^{(2)})$:

- A term conveying some statistics, which entails spatial regularization. To interpolate the missing data, the work in [10] enforced a histogram prior using the Wasserstein distance, linking the approach with the optimal transport [10]. The originality of this technique consists of the ability to consider multi-histogram priors without being constrained to a parametric model. In this project, we aim at learning the statistical models of the medical images to infer them into the Wasserstein distance.

- A term enforcing a new type of nonlocal regularization, which involves statistical information that is not spatially indexed. One can learn a nonlocal graph [10, 11] that provides better connections between the missing pixel, by taking into account the rich information modeled by graphs in the high-resolution snapshots.

- Deep plug and play super-resolution, which consists of integrating convolution neural network-based image denoisers [12] as a regularization term in the optimization scheme [13].

3 Goals and objectives of the project

In this proposal, we consider the following objective:

- Design adequate regularization compatible with the texture images.

- Efficient algorithm to treat the whole slide in an efficient manner

4 Significance of the project

5 Work plan

The research plan covers the work of the applicant (i.e., P. Frossard, 10%) - no financial support required), as well as one Postdoctoral researcher (i.e., M. El Gheche, 50%) for whom we request financial support for one year. The work plan is presented below: (Ajouter un tableau)

This project should lead to the development of a large-scale FNS grant proposal in the area of representation learning for graphs, where the definition of a metric to compare signals is of paramount importance.

References

[1] K. Zhang, X. Gao, D. Tao, and X. Li, "Single image super-resolution with non-local means and steering kernel regression," *IEEE Trans. Image Process.*, vol. 21, no. 11, pp. 4544–4556, Nov. 2012.

[2] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1-4, pp. 259–268, Nov. 1992.

[3] G. Gilboa and S. Osher, "Nonlocal operators with applications to image processing," *Multiscale Modeling and Simulation*, vol. 7, no. 3, pp. 1005–1028, Nov. 2009.

[4] D. Dai, H. Riemenschneider, and L. V. Gool, "The synthesizability of texture examples," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2014, pp. 3027–3034.

[5] G. Cross and A. Jain, "Markov random field texture models," *IEEE Trans. Pattern Anal. Mach. Int.*, no. 1, pp. 25–39, Jan. 1983.

[6] A. Kaspar, B. Neubert, D. Lischinski, M. Pauly, and J. Kopf, "Self tuning texture optimization," *Comput. Graph. Forum*, vol. 34, no. 2, pp. 349–359, 2015.

[7] J. Portilla and E. Simoncelli, "Image denoising via adjustment of wavelet coefficient magnitude correlation," in *Proc. IEEE Int. Conf. Image Process.*, vol. 3, Vancouver, Canada, Sep. 2000, pp. 277–280.

[8] B. Galerne, Y. Gousseau, and J.-M. Morel, "Random phase textures: Theory and synthesis," *Image Process On Line*, vol. 1, 2011.

[9] G. Tartavel, Y. Gousseau, and G. Peyré, “Variational texture synthesis with sparsity and spectrum constraints,” *J. Math. Imaging Vision*, vol. 52, no. 1, pp. 124–144, May 2015.

[10] M. El Gheche, J. Aujol, Y. Berthoumieu, and C. Deledalle, “Texture reconstruction guided by a high-resolution patch,” *IEEE Transactions on Image Processing*, vol. 26, no. 2, pp. 549–560, Feb 2017.

[11] G. Chierchia, M. El Gheche, G. Scarpa, and L. Verdoliva, “Multitemporal sar image despeckling based on block-matching and collaborative filtering,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 10, Oct 2017.

[12] S. Arjomand Bigdeli, M. Zwicker, P. Favaro, and M. Jin, “Deep mean-shift priors for image restoration,” in *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds. Curran Associates, Inc., 2017, pp. 763–772. [Online]. Available: http://papers.nips.cc/paper/6678-deep-mean-shift-priors-for-image-restoration.pdf

[13] J. Jiang, Y. Yu, J. Hu, S. Tang, and J. Ma, “Deep CNN denoiser and multi-layer neighbor component embedding for face hallucination,” in *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden.*, 2018, pp. 771–778. [Online]. Available: https://doi.org/10.24963/ijcai.2018/107